

УДК 81-13

DOI 10.21661/r-554295

Ивлиева И.В.

**ПОТЕНЦИАЛЬНЫЕ ЕДИНИЦЫ ГЛАГОЛЬНОГО СИНТЕЗА:
СВОЙСТВА И ВОЗМОЖНОСТИ (НА МАТЕРИАЛЕ
РУССКИХ ГЛАГОЛОВ ЗВУЧАНИЯ)**

***Аннотация:** цель исследования – выявить особенности потенциальных модификаций и лакунообразующих модификаторов, участвующих в процессе синтеза глагольного значения. Научная новизна исследования заключается в дифференцированном представлении лексико-семантических и морфологических средств (глагольных модификаторов), способных, с одной стороны, синтезировать заданное значение (глагольную модификацию), а с другой, привести к образованию лакун. В результате анализа конкретных модификаторов и модификаций группы глаголов звучания выявлены новые элементы синтеза модификаций (потенциальные модификации), показаны особенности появления лакун с помощью лакунообразующих модификаторов. С привлечением сведений Национального корпуса русского языка и специального проекта научно-технологического университета штата Миссури, США, впервые в лексикографическую практику введены модификации, ранее считавшиеся несуществующими.*

***Ключевые слова:** синтез русского глагола, словообразовательный синтез, глаголы звучания, потенциальная лексика, узуальная лексика, лакуна, лакунообразующий модификатор, глагольный модификатор, глагольная модификация, Национальный корпус русского языка.*

Введение

Современный этап исследования лексики, полученной в результате синтеза нового значения, характеризуется рядом особенностей. С одной стороны, довольно подробно и полно описаны механизмы появления модификаций в группе глаголов звучания, выявлен список возможных модификаторов, участвующих в появлении нового слова [5, с. 81], существует описание модифицированности на

корневом уровне [5, с. 112–115], разработано словарное описание узусальной лексики, полученной в процессе синтеза [6].

С другой стороны, новые сведения о свойствах синтезированных дериватов, устройстве модификационных гнезд и появление лакунарной лексики неизбежно ставят перед исследователями другие, более сложные вопросы: каковы причины присоединения / неприсоединения того или иного модификатора к исходному глаголу, как определяется соотношение между фактическими, кодифицированными единицами синтеза и потенциально возможными, каковы правила заполнения глагольных модификационных матриц и многие другие. Решение этих вопросов по-прежнему направлено на главную задачу синтеза: «объяснить механизм порождения и понимания новых слов» [18].

И.С. Торопцев, изучая механизмы словопроизводственных моделей, справедливо отмечал сложность и неординарность процесса синтеза: «Стремясь получить словообразовательным синтезом новые значения путем комбинации морфем, мы тем самым допускаем, что все лексические значения слов, которые когда-либо потребуются языку, в нем уже есть, только в разобранном виде. Это в принципе неверно. <...> Источник нашего знания в объективной действительности, а не в морфемах» [17].

Изучение механизма словопроизводства в синтезе начинается с построения звуковой оболочки слова (или иной семантической единицы, модификации) и углубляется в экстралингвистические факторы. Исследованию модификаций глагольной лексики, как узусальных, так и потенциальных, предшествует работа ученых во многих областях лингвистики.

В отечественной науке создана серьезная теоретическая база, на основе которой стали возможны исследования в области синтеза. Первые попытки синтезировать слово на базе имен существительных и моделировать механизм появления деривата описаны в трудах И.Г. Милославского [12]. Работы Е.А. Земской, М.А. Кронгауза легли в основу изучения семантики приставок и сочетаемости их с основами глагола [2; 9]. На базе исследований Е.С. Кубряковой, Л.Г. Бабенко, Н.А. Янко-Триницкой и др. по семантическому анализу глагола стало

возможным изучение дополнительных смысловых приращений [10; 1; 19]. Труды И.С. Улуканова и А.Н. Тихонова [18; 16] позволили расширить понимание узуальной и потенциальной лексики и выйти на уровень представления результатов синтеза в лексикографии.

В лексикографической традиции слово считается полноценной языковой единицей только в том случае, «если оно встречается в обследованных текстах хотя бы два раза» [8, с. 102]. Потому так важно извлекать новый материал из различных баз и источников, находить новые единицы и вводить их в научный оборот. В этой связи одной из важнейших задач синтеза значений считаем не только выявление и семантическое описание полученных единиц, но и лексикографическую разработку потенциальной лексики, то есть каждая единица синтеза в перспективе должна быть кодифицирована.

Актуальность выбранной темы заключается в исследовании наименее известного участка модификационной системы глаголов звучания – потенциальной лексики. Без семантического описания потенциальных единиц синтеза, установления этапов их лексикализации, картина появления нового слова представляется незавершенной.

Достижению *цели* исследования способствует решение следующих *задач*:

- 1). Определить систему потенциальных модификаций и лакунообразующих модификаторов глаголов звучания;
- 2). Провести анализ значений потенциальных единиц на уровне ЛСВ и выявить новые единицы синтеза – потенциальные модификации;
- 3). С помощью методики веб-извлечения (web scraping) глагольных парадигм установить модификации, не учтенные современной лексикографией и ранее считавшиеся несуществующими.

Методы исследования, которые использовались для выявления лексических средств, участвующих в формировании нового значения, следующие: сравнительно-сопоставительный метод лексических значений, метод сплошной выборки лексических единиц со значениями различных модификаторов, метод компонентного анализа значения, комбинаторно-семантический анализ контекста,

метод веб-извлечения (web scraping) русских глагольных парадигм из электронных словарей и баз данных.

Практическая значимость исследования заключается в том, что материал, представленный в статье, может быть использован как лингвистами, изучающими процессы синтеза нового значения, так и преподавателями русского языка как иностранного при изучении спецкурсов по морфологии и словообразованию русского языка. Материал также полезен для спецсеминаров по лексикографии.

Система потенциальных модификаций и лакунообразующих модификаторов глаголов звучания

Глагольные модификации, которые по разным причинам отсутствовали в лексикографических источниках, в предыдущих исследованиях (в период с 2003 по 2020 год) рассматривались нами в категории лакунарных (как полностью отсутствующие элементы) [5; 6; 7]. В исследованиях других авторов, рассматривающих синтезированное значение, данная категория слов также носит название *лакуны* или рассматривается в составе окказиональной лексики [18].

В настоящей статье в составе лакун проведена более детальная дифференциация: выделены единицы узуальной, потенциальной лексики и собственно лакун. То есть группа лакун в данном исследовании получает более углубленное описание: из системы лакун корректно выведена группа модификаций, которая имеет свойства потенциальной и узуальной лексики.

В самом названии «потенциальная единица» заложен смысл, отличающий данные слова от узуальной и лакунарной лексики. И рассмотрение потенциальной лексики в составе лакун имело под собой основание – ее некодифицированность. Это одно из свойств потенциальных единиц. Но при этом не учитывалось ее другое свойство – встречаемость в речи и функционирование в текстах. Новый статус данных единиц дает возможность взглянуть на не существующую на данный момент модификацию (не представленную в словарях) как на потенциальную единицу синтеза. Единицу, которая со временем будет кодифицирована, так как имеет на то основания. Рассмотрим эти основания и процесс выявления потенциальных модификаций.

Ранее нами установлено, что синтез значения может развиваться в четырех направлениях: 1) модификатор присоединяется к исходному значению и образуется дериват с заданным значением; 2) модификатор присоединяется к исходному значению, но в результате появляется один или несколько дополнительных смыслов; 3) модификатор присоединяется к исходному значению, но результат синтеза не зафиксирован словарями; 4) и наконец, направление отрицательного синтеза, когда соединению препятствуют различные факторы лингвистического или экстралингвистического характера [7].

В результате синтеза *первого типа* появляется УЗУАЛЬНОЕ значение; обеспечивается предсказуемый результат и с помощью того или иного модификатора появляется новый дериват: *запеть* (от *петь* + *за-*), *выпеть* (от *петь* + *вы-*), *раскричатся* (от *кричать* + *рас-*ся-) и т. п. (Список модификаторов и их подробное описание даны в «Экспериментальном модификационном словаре русского языка» [6]).

Сюда же мы относим слова, которые фиксируются словарями только в одном или нескольких своих значениях (орфографические, грамматические, аспектные и т. п.).

В типах синтеза 2, 3 и 4 наблюдается появление лексем, которые не невозможно было спрогнозировать в соответствии с основной формулой синтеза: $X_{ит} + Y_{1...30} = Z_n$, где $X_{ит}$ – исходный глагол, $Y_{1...30}$ – модификатор из тридцати возможных, Z_n – результат синтеза, который, в свою очередь, может иметь несколько вариантов.

Вариант 1. *Появление модификаций с приращенными смыслами*, которые отсутствуют в исходном глаголе и в значении модификатора. Это так называемые *индивидуальные модификации*, которые подробно описаны в модификационном словаре [6].

Например, в индивидуальной модификации *дочихать* (чихать + *до-*) в конечном значении «кончить, перестать чихать» появляется сема финитивности вместо ожидаемого значения «до некоторой точки», который должен передаваться модификатором *до-*, а в значении *отпеть* (петь + *от-*) вместо семы

финитивности «кончить петь», появляются семантические приращения «в церкви», «обряд», «усопший» и т. п.

При этом в одном из своих лексических вариантов (ЛСВ), в значении *отпеть* «кончить петь», может сохраняться и исход синтеза по формуле $X_{иг} + Y_{1...30} = Zn$: *Петухи отпели* («кончили петь»), *рассвело*.

Такие модификации находятся на границе между группами узуальной и потенциальной лексики. Формально они узуальные, так как представлены в системе русского языка, но с точки зрения звуковой семантики, модификации типа *отпеть* и *дочихать* («кончить петь» и «кончить чихать») являются потенциальными, так как их появление в большей степени обусловлено контекстом и коммуникативной необходимостью.

Вариант 2. *Появление условных лакун или потенциальных модификаций*. Это модификации, которые объективно трудно установить, так как о них нет информации в лексикографических источниках. Лексемы, подобные *раскукарекаться*, *разбарабаниться* и т. п. формально не существуют. Но относить их к разряду лакун тоже не совсем корректно. Наше исследование показывает, что подобные потенциальные модификации следует вывести из списка лакун по нескольким причинам, о которых будет сказано ниже.

Вариант 3. *Собственно лакуны*. Это единицы с несочетаемыми морфемами, то есть семантика исходного глагола и семантика присоединяемых аффиксов противоречат друг другу по тем или иным причинам. Так, например, неприсоединяемыми смыслами в ряде глаголов звучания являются сочетания предельных глаголов с модификаторами мгновенности *-ну-* и *вз/вс-ну-* {Mod4 мгновенно + один раз}: *петь* + *-ну-* (**певнуть*), *говорить* + *-ну-* (**говорнуть*), *басить* + *-ну-* (**баснуть*), *картавить* + *-ну-* (**картавнуть*), *сопеть* + *-ну-* (**сопнуть*) и т. п. Хотя неопредельные глаголы активно образуют модификации с данным смыслом. Ср., например, *квакать* – *квакнуть*, *кричать* – *крикнуть*, *чихать* – *чихнуть*, *мяукать* – *мяукнуть*, *скрипеть* – *скрипнуть* и т. д.

И если свойства узуальных и лакунарных модификаций более или менее ясны (узуальные модификации предсказуемы и активно синтезируются в

результате соединения модификаторов с исходными глаголами; лакунарные модификации предсказуемы в противоположных процессах – в несочетаемости глаголов с теми или иными модификаторами), то свойства потенциальных единиц мало исследованы и могут показывать самые противоречивые результаты. С одной стороны, они могут оказаться в составе модификаций-лакун, с другой – со всей очевидностью выступать в роли узуальных единиц.

Рассмотрим, как строится процесс вывода потенциальных единиц в зону узуальных (кодифицированных) лексем на примере модификации *погремливаться*. Данная модификация не представлена ни одним словарем русского языка, и соответственно, не имеет толкования. Но с помощью поисковой программы [20; 21] в НКРЯ находим пример использования в литературном тексте: *На небе со всех сторон надвигались черные, как вороново крыло, тучи, и начинало уж вдали погремливаться* (А.Ф. Писемский. Старческий грех) [13].

Глагол *погремливаться*, как и аналогичные ему, теоретически находится в составе лакун русского языка, поскольку не разработан ни одним словарем. Но фактически является положительным результатом синтеза, поскольку образован путем присоединения трехкомпонентного модификатора *по-и-ва-* {некоторое время + много раз + слабо} к исходному глаголу *греметь*. Синтезированное значение при этом появляется в контексте с прогнозируемым смыслом. Ср., например, значения синонимов *погрохатывать* и *погромыхивать*:

Погрохатывать. Производить грохот, шум в течение некоторого времени [15].

Погромыхивать. Время от времени слегка громыхать [15].

Следовательно, значение *погремливаться* можно представить в виде аналогичного толкования:

Погремливаться. Греть слегка или время от времени.

Для того чтобы протестировать сочетаемость полученной модификации, помещаем ее в один контекст с синонимами: *На небе со всех сторон надвигались черные, как вороново крыло, тучи, и начинало уж вдали погремливаться* (ср.: *погромыхивать, погрохатывать* – И.И.) (А. Ф. Писемский. Старческий грех) [13].

В приведенном «звуковом» контексте, как видим, все модификации, как узуальные (*погромыхивать* и *погрохатывать*), так и потенциальная (*погремливать*), адекватно передают значение модификатора *по-ыва-/-ива-*.

Создавая значение для потенциальной единицы, мы тем самым выводим ее из состава лакун и переводим в разряд узуальной лексики, увеличивая модификационный потенциал глагола *греметь* и активность модификатора Mod6 *по-ива-* {некоторое время + много раз + слабо}.

Таким образом, можно сделать вывод, что основным свойством потенциальных единиц, которые мы планируем вывести из состава лакун, является их прогнозируемость и подобность узуальным единицам, их принципиальная кодифицируемость. Результаты данного исследования дают возможность многим потенциальным модификациям изменить свой статус с лакунарных единиц на узуальные.

Другой единицей синтеза, которая влияет на формирование состава потенциальной лексики, являются *лакунообразующий модификатор*. Исключив из состава модификаций полные лакуны, выделенные с помощью лакунообразующих модификаторов, мы тем самым сузим круг потенциальной лексики и максимально уточним его.

Как следует из названия модификатора, данные единицы всегда прогнозируют появление лакун. Главным вопросом здесь является наличие информации о том, какой именно модификатор и для какого конкретного глагола является лакунообразующим. Так как один и тот же модификатор для разных глаголов может выступать и в роли лакунообразующего, и в роли синтезирующего положительный результат. Выше это было показано на примере модификаторов мгновенности *-ну-* и *вз/вс-ну-*, где: с одной стороны, их присоединение ведет к образованию лакун типа **невнуть*, **картавнуть*, с другой, данные модификаторы активно участвуют в синтезе узуальных единиц типа *крикнуть*, *свистнуть*, *мяукнуть* и т. п. Так что грань между лакунами и потенциальными модификациями довольно размытая.

Мы классифицируем лакуны по двум типам: 1) по наличию / отсутствию в их значениях звуковой семантики и 2) по количеству лакунарных значений в лексеме.

К *первому типу* относятся лакуны, которые по своей семантике могут отличаться отнесенностью к звуковому ряду: это так называемые *лакуны-значения* (терминология наша – И.И.), содержащие «значение звука», и *лакуны-формы* (терминология наша – И.И.), в которых сема звука отсутствует. Так происходит, например, в модификации *наплакать* в контексте *кот наплакал* (в значении «мало»), где форма модификации присутствует в системе (плакать + *на-* {некоторое время + один раз + слабо}), но переноса звукового значения в синтезе не происходит.

В первом случае (лакуна-значение / лакуна-форма) дериват *наплакать* существует в языке, но как звуковая модификация отсутствует. Поэтому трехкомпонентный модификатор *на-* {некоторое время + однократно + слабо} = *наплакать* для глаголов типа *плакать* является лакунообразующим. Хотя для других глаголов звучания он вполне актуален: *наговорить* (говорить + *на-*), *нажужжать* (жужжать + *на-*), *назвонить* (звонить + *на-*), *наиграть* (играть + *на-*), *насвистеть* (свистеть + *на-*), *нашептать* (шептать + *на-*) и т. п.

Здесь важно отметить следующую закономерность, которую мы выводим из анализа модификаций типа *лакуна-значение / лакуна-форма*. Если форма модификации существует хотя бы в одном из значений производного глагола (даже не «звуковом»), то форму следует отнести не к лакунам глаголов звучания, а к потенциальным модификациям. Дело в том, что мы не имеем возможности заранее предусмотреть все контексты, в которых может появиться в перспективе данная модификация (в том числе окказионально). Окказиональность употребления модификации непредсказуема, потому она не исключает и появления звуковых смыслов.

Ко второму типу лакун относим *лакуны полные* и *лакуны ЛСВ*. Они отличаются по количеству лакунарных значений в итоговой модификации, когда лакунарность передается либо всеми значениями слова, либо только одним из них.

Лакуны полные и *лакуны ЛСВ* появляются в результате попытки присоединить к исходному глаголу не присоединяемый (лакунообразующий) модификатор; при этом в синтезе участвуют один или несколько вариантов значения. Примером полных лакун являются модификации с неприсоединяемыми смыслами: **вскартавнуть* (картавить + *вс-ну-*), **встукнуть* (стучать + *вс-ну-*), **вспевнуть* (петь + *вс-ну-*), **всшепелявнуть* (шепелявить + *вс-ну-*) и т. п. Ни в одном из своих значений ЛСВ глаголы *картавить*, *стучать*, *петь* и *шепелявить* не способны присоединять смысл трехкомпонентного модификатора *вз-/вс- ...-ну-/ -ану-* {начало + мгновенно + один раз}.

Лакуны-ЛСВ имеют сходство с *лакунами-формами*, когда модификация в одном или нескольких значениях в процессе синтеза «теряет» звуковые семы: *прозевать* в смысле «упустить благоприятный момент» (*прозевать поезд, прозевать пешку*) [13]; *простучать* «начать часто и сильно пульсировать» (*простучало в висках, в голове*) [13] и т. п.

Каждый модификатор, в зависимости от исходного глагола, может быть как активным (приводить к образованию узуальной лексики), так и нестандартным (приводить к появлению индивидуальных модификаций) и лакунообразующим (приводить к появлению лакун). И у каждого глагола свой рейтинг активных, индивидуальных и лакунообразующих модификаторов.

Так, например, у глагола *чихать* выделяем следующие типы модификаторов, ведущие к образованию разных пластов лексики:

Активные модификаторы (узуальная лексика): до-ся (*дочихаться*); за- (*зачихать*); на-ся (*начихаться*); от- (*отчихать*); по- (*почихать*) и т. п.

Индивидуальные модификаторы (потенциальная лексика с переносом звуковой семы): до- (*дочихать* «кончить, перестать чихать»);

Модификаторы, ведущие к лакунам-ЛСВ (потенциальная лексика без переноса звуковой семы): на- (*начихать* «отнестись с пренебрежением»);

Лакунообразующие модификаторы (лакуны с невозможностью присоединения звуковых сем): при-ну- (**причихнуть*); с- (**счихать*) и т. п.

Анализ значений потенциальных единиц на уровне ЛСВ

Впервые тема анализа значений модификаций на уровне ЛСВ была затронута в 2003 году в статье «Особенности синтеза многозначных глаголов. Семантическое варьирование морфем» [3]. Тогда была предпринята попытка промаркировать ЛСВ некоторых глаголов звучания и выяснить, насколько различные варианты значения одного и того же глагола способны участвовать в синтезе, и какие результаты дает подобное семантическое расщепление.

На материале одной лексической группы было установлено, что у многозначных глаголов разные части значений по-разному присоединяют одни и те же модификаторы. Так, например, при присоединении к исходному глаголу одного и того же модификатора на выходе синтеза можем получить модификации с несколькими значениями: одно прогнозируемое, а другие с дополнительными приращенными смыслами:

Значение (А) «начать говорить': говорить + за- {начало + один раз} (*заговорить* «начать говорить'); *Хотел заговорить с ней, но о чем заговорить, не знал и оттого не заговорил* (В. Астафьев. Пролетный гусь) [13].

Значение (Б) «воздействовать на кого-л. заговором, заклинанием': говорить + за- {начало + один раз} (*заговорить* «воздействовать на кого-л. заговором, заклинанием'); *Через полчаса можно будет заговорить рану, но пока слишком мало сил* (С. Лукьяненко. Ночной дозор) [13].

Значение (В) «утомить продолжительным разговором': говорить + за- {начало + один раз} (*заговорить* «утомить продолжительным разговором'); *Он прекрасно умел вести беседу, мог заговорить любого* (А. Маринина. Стечение обстоятельств) [13].

В приведенных выше примерах приращенные значения модификаций (Б) «воздействовать на кого-л. заговором, заклинанием' и модификации (В) «утомить продолжительным разговором' не представлены ни в значении исходного глагола *говорить*, ни в значении модификатора Mod2 *за-*, данные смыслы являются непрогнозируемыми, он появляются у деривата, как у самостоятельной единицы, связанной только частично со значением исходного глагола.

Для того чтобы учесть все возможные контекстные модификации глагола, и учесть их в дальнейших исследованиях, необходимо дифференцировать их и ввести обозначения для различных модификаций на уровне ЛСВ. При этом мы получим максимально верные сведения о количестве синтезированных лексем, что в свою очередь даст информацию о потенциальных «звуковых» модификациях. В дальнейшем данная информация может быть использована при составлении словаря потенциальной лексики и лакун.

Следует учитывать также и влияние переносных значений на результат синтеза: некоторые из них сохраняют звуковую сему, а другие нет (*провыл пес, провыл человек, ветер провыл*, но: *ноет ребенок и ноет зуб*). Сегодня можно с уверенностью сказать, что результат модификации напрямую зависит от того, какое именно значение ЛСВ мы учитываем.

Рассмотрим, как ведется расчет значений лакун, и в частности, *лакун-ЛСВ*. У *лакун-ЛСВ* может быть неопределенное количество значений, допустим – n , при этом n всегда больше 1 ($n > 1$), но меньше общего количества ЛСВ слова ($n < P$).

Пример маркировки лакун одной лексемы и одной модификации (глагола *стучать* и модификации *настучать*) выглядит следующим образом:

Стучать – *настучать* (стучать + на- {некот. время + однократно + слабо}):

Лакуна 1 / ЛСВ 1. Стучать в дверь (**настучать в дверь*);

Лакуна 2 / ЛСВ 2. Стучать на машинке (*настучать на машинке* – «напечатать»); *Я приехала в прокуратуру и быстро настучала телеграмму от своего имени* (Е. Топильская. Тайны реального следствия) [Там же].

Лакуна 3 / ЛСВ 3. Стучать по крыше (капли дождя) (*дождь *настучал*);

Лакуна 4 / ЛСВ 4. Стучать на товарищей (*настучать на друга* – «донести, наябедничать»). *Некоторым легче настучать, чем постучать в стенку* (В. Войнович. Монументальная пропаганда) [Там же].

В группе глаголов «поведения» модификация *настучать* скорее всего будет отвечать всем правилам синтеза, но в группе звуковых глаголов этот вариант ЛСВ 4 является лакуной, так как не поддерживает перенос семы звука.

Лакуна 5 / ЛСВ 5. Стучать по голове (настучать по голове «обругать». Тебе чего, по голове настучать? Быстренько проваливай, пока цел (С. Романов. Парламент) [Там же].

*Лакуна 6 / ЛСВ 6. Стучать на барабане (*Настучать на барабане «наиграть»).*

Таким образом, у глагола *стучать* в звуковом значении при соединении с модификатором *на-* в результате синтеза образуется 6 лакун (*настучать_{n=6}*) и ни одной узуальной модификации со значением *стучать* + *на-* {некоторое время + однократно + слабо}. Потенциальными и узуальными можно рассматривать вариант *настучать мелодию*, если будет найден контекст, подтверждающий данный смысл. Отсутствие в системе единицы *настучать* (в звуковом значении) говорит не об отсутствии самого денотата или концепта в экстралингвистической действительности, а скорее о ее коммуникативной не востребованности на данный момент.

Кроме того, исследование показывает, что значение ослабленности и незначительной силы однократного действия, выраженного приставкой *на-* {некоторое время + однократно + слабо} присоединяется лишь к незначительной части глаголов звучания: *набормотать, наворчать, наговорить, нагреметь, нажужжать, наиграть, накаркать, накуковать, напеть, (нарычать – индивидуальная модификация – не зафиксирована в словарях), насвистать, насвистеть, наскрипеть, наскулить, насмеяться, настукать, настучать (незвуковое значение), натопать, натрезвонить, натрещать, натрубить («много, долго потрубить» – индивидуальное значение), нафыркать, нахлопать, начиркать («чиркая, испортить» – не звуковое значение), начихать, нашаркать, нашептать, нашмыгать, нашуметь, нащелкать* [14, с. 631–675].

И данная тенденция модификаторов – не соединяться со значениями ЛСВ определенной семантики (как правило, узкой) наблюдается у многих модификаторов, что в конечном итоге ведет к лакунарности.

Метод веб-извлечения (web scraping) глагольных парадигм. Выявление потенциальных («несуществующих») модификаций

Метод веб-извлечения (web scraping) глагольных парадигм разработан автором совместно с Перри Кооб [Koob, P., Ivliyeva, I.] исключительно для поиска синтезированных значений в существующих электронных базах данных.

Сразу отметим, что «несуществующие» модификации – это условное название единиц, которые на самом деле могут быть вполне реальными, и отличаться от узуальной лексики лишь отсутствием фиксации в словарях. В начале исследования мы описали систему потенциальных модификаций, показали классификацию и свойства данных единиц. Далее представим методику выявления потенциальных единиц, которая может быть применена в дальнейшем к любой семантической группе.

Исследование основано на коллекции текстов Национального корпуса русского языка, представляющей современный русский язык периода 1950–2007 годов. Объем цитатного материала, на котором строится поиск модификаций, составляет 92 млн. словоупотреблений. «В настоящее время Национальный корпус включает электронные тексты письменного русского языка XVIII – начала XXI века (основной корпус), записи устной речи, газетный, поэтический и диалектный корпусы, < ... > корпус древнерусского языка» [11]. Таким образом, можно утверждать, что поиск модификаций ведется в контекстах всех жанров.

Алгоритм поиска выглядит следующим образом:

1. На первом этапе с помощью метода веб-извлечения (web scraping) [20; 21] выявляем искомые модификации, исключив из рассмотрения только узуальную лексику. На основе табличных данных составляем корпус *потенциальной* и *лакунарной* лексики;

2. Второй этап поиска – это *дифференциация потенциальных модификаций* и *лакун*; заполняем *поглагольные таблицы*, маркируя лакуны и потенциальные модификации;

3. На третьем этапе потенциальные единицы *проверяем по НКРЯ* и другим Интернет источникам, собирая доказательные примеры употребления;

4. Выводим уточненный *список новой лексики* с целью ее дальнейшей кодификации, переводя тем самым лексемы из разряда потенциальной лексики в разряд узуальной.

Остановимся подробнее на каждом из этапов.

Для составления первоначального списка потенциальной и лакунарной лексики следует обозначить варианты, которые не вызывают сомнений с точки зрения их узуальности: количество их словоупотреблений в НКРЯ варьируется от десятков единиц (*выбивала* – 69 употреблений) до нескольких тысяч (*пробить* – 2277 употреблений).

Ниже в *Табл. 1* узуальные единицы выделены курсивом; они показаны в четырех формах (инфинитив, прошедшее время: жен.р., муж.р. и ср.р.), к каждой из которых дано количество их употреблений в НКРЯ. Поисковые таблицы здесь и далее приведены в сокращенном виде, адаптированном для формата статьи. Полные варианты таблиц представлены на сайте *Web scraping project by Missouri S&T, IT and ALP departments* [20].

Таблица 1

Отбор узуальной лексики из списка модификаций всех типов

Дериват / модификация	Прош. жен.р.	Прош. муж.р.	Прош. ср.р.	Кол-во словоупотреблений в НКРЯ			
отбахать	отбахала	отбахал	отбахало	NA	NA	NA	NA
побахать	побахала	побахал	побахало	NA	NA	NA	NA
пробахать	пробахала	пробахал	пробахало	NA	NA	NA	NA
разбахаться	разбахалась	разбахался	разбахалось	NA	NA	NA	NA
<i>выбивать</i>	<i>выбивала</i>	<i>выбивал</i>	<i>выбивало</i>	<i>331</i>	<i>69</i>	<i>169</i>	<i>232</i>
<i>забить</i>	<i>забила</i>	<i>забил</i>	<i>забило</i>	<i>450</i>	<i>231</i>	<i>586</i>	<i>570</i>
<i>пробить</i>	<i>пробила</i>	<i>отбил</i>	<i>пробило</i>	<i>658</i>	<i>324</i>	<i>726</i>	<i>569</i>
<i>отбить</i>	<i>отбила</i>	<i>пробил</i>	<i>отбило</i>	<i>667</i>	<i>211</i>	<i>575</i>	<i>618</i>

В следующую таблицу (Табл. 2) помещаем *потенциальную* лексику и *лакуны-ЛСВ*. Анализ цитатного материала из НКРЯ проводим для каждой модификации, вручную снимая варианты омонимии.

Так, например, модификация *проахал* представлена в НКРЯ в «незвучковом» значении «потерять, лишиться чего-л., благодаря бездействию». В глаголах

звучания данную модификацию следует рассматривать как лауну-ЛСВ, так как существует еще и звуковое значение глагола *проахать*.

Ср., например, употребление в текстах лауны-ЛСВ и модификации *проахла* в звуковом значении: *Дальше Шкуропатенко <...> бельмом уставился в окошко, доглядает <...>. Толь-то проахал, дядя. Бухгалтера два, тоже ээки, хлеб поджаривают на печке* (А. Солженицын. Один день Ивана Денисовича) [13]. *Ах, мне бы такое платье, когда я была маленькой... - восхищенно проахла мама* (И. Шиповских, Сказка о пуговичках) [Там же].

В табл.2 модификации лауны-ЛСВ выделены кавычками.

Таблица 2

Корпус потенциальной лексики и лакун-ЛСВ в НКРЯ (со снятой омонимией)

Дериват / модификация	Представленность словоупотреблений в НКРЯ				
	дериват -ть/ -ться	жен.р. -ла/ -лась	муж.р. -л / -лся	ср.р. -ло / -лось	мн.ч. -ли / -лись
проахать	-	-	«проахал»	-	-
отбренчать	-	-	отбренчал	-	-
побренькать	-	-	побренькал	-	-
поверещать	-	-	поверещал	-	-
разверещаться	-	-	-	-	разверещались
повопить	повопить	повопила	-	-	-
наворчаться	-	-	наворчался	-	-
навыться	навыться	-	навылся	-	-
прогалдеть	-	-	прогалдел	-	прогалдели
прогаркать	-	-	прогаркал	-	-
прогикать	-	-	прогикал	-	-
разгоготаться	разгоготаться	-	разгоготался	-	-

Таким образом, из примеров, содержащихся в НКРЯ, методом сплошной выборки извлекаем и группируем по лексическому составу модификаций образцы контекстов. Цитатный материал дает нам достоверную информацию о дериватах, существование которых изначально можно было прогнозировать лишь интуитивно.

Далее, используя критерии и информацию о лакунообразующих модификаторах и данные поисковых таблиц [20], заполняем *поглагольные матрицы* с указанием всех модификаций. На данном этапе происходит окончательная дифференциация потенциальных, узуальных и лакунарных модификаций. Покажем вариант заполнения матрицы на примере глагола *выть* и его дериватов (Табл. 3).

Таблица 3

Модификационная матрица глагола **ВЫТЬ**

M1 <i>по-, про1 повыть, провыть</i>	M2 <i>за- завыть</i>	M3 <i>про2, *с- провыть *свыть</i>	M4 <i>-ну- LACUNA</i>	M5 <i>раз-ся развиться</i>	M6 <i>по-ыва- повывать</i>
M7 <i>до-ся довыться</i>	M8 <i>про-ся, *вы- ся, на-ся провыться, навыться</i>	M9 <i>от- отвыть</i>	M10 <i>при-ыва- LACUNA</i>	M11 <i>на-ыва- навывать</i>	M12 <i>из-ся, об-ся извыться, обвыться</i>
M13 <i>при-ну-, под- подвыть</i>	M14 <i>вз- взвыть</i>	M15 <i>вз-ыва- взвывать</i>	M16 <i>*от-ыва- POTENTIAL</i>	M17 <i>про-ыва- POTENTIAL</i>	M18 <i>недо- POTENTIAL</i>
M19 <i>вз-ну- LACUNA</i>	M20 <i>*вы-ыва- LACUNA</i>	M21 <i>вы- LACUNA</i>	M22 <i>пере- POTENTIAL</i>	M23 <i>пере-ыва- POTENTIAL</i>	M24 <i>под-ыва- подвывать</i>
M25 <i>за-ыва- завывать</i>	M26 <i>на- POTENTIAL</i>	M27 <i>недо-ыва- POTENTIAL</i>	M28 <i>до- POTENTIAL</i>	M29 <i>до-ыва- POTENTIAL</i>	M30 <i>-ва- LACUNA</i>

Поглагольная матрица наглядно демонстрирует модификации всех типов:

Узуальные: M1 *повыть, провыть*; M2 *завыть*; M3 *провыть*; M5 *развиться*; M6 *повывать*; M7 *довыться*; M8 *навыться* и т. п.

Потенциальные (в том числе условно возможные варианты и окказиональные: *перевыть, перевывать, навыйть, довыть, довывать* и т. п.

Лакуны полные: труднопроизносимые и неидентифицируемые для носителей языка сочетания морфем типа *выть* + *-ну-* = **вы(в)нуть*; *выть* + *вы-* = **вывыть*; *выть* + *вз-ну-* = **взвы(в)нуть* и т. п.

На данном этапе исследования критерием отнесения модификации к тому или иному типу служат примеры употребления модификаций в НКРЯ (их наличие или отсутствие в базе) и непротиворечивые контексты со снятой омонимией.

У каждого глагола при этом свой набор модификаций и своя степень заполнения матрицы. У глаголов более узкой семантики процент лакун в матрице будет больше, а потенциальных единиц меньше.

Необходимо отметить, что автоматический поиск ведется по тем спискам, которые составлены изначально вручную в 2013 году на базе толковых и аспектных словарей. Корректировка списков с учетом новых данных поисковых программ и маркировки значений ЛСВ входит в одну из перспективных задач будущих исследований.

Следующим этапом выявления потенциальной лексики глаголов звучания является поиск цитатного материала с участием потенциальных единиц не только по базе НКРЯ, но и по другим Интернет-источникам.

Рассмотрим данный этап на примере исходного глагола *бабахать*. Модификационное гнездо данного глагола, по данным современных словарей и других источников, состоит всего из двух единиц: *бабахнуть* и *забабахать*.

Бабахнуть. 1. Издать сильный отрывистый звук или раздаться (о звуке); 2. С силой ударить, стукнуть, выстрелить [15]. *Где-то далеко бабахнуло, – перечисляет Аня. – Гудит высоковольтная* (Татьяна Соломатина. Большая собака, или «Эклектичная живописная вавилонская повесть о зарытом») [13].

Забабахать. Начать бабахать [15]. *А под утро они как налетели, как забабыхали!* (Г. М. Марков. Строговы) [13].

Других лексикографически зафиксированных результатов синтеза в русских словарях не представлено. В модификационном словаре русского языка также разрабатываются только две модификации: Mod4 *-ну-* {мгновенно + один раз} и Mod2 *за-* {начало + один раз}.

Используя данные поисковых таблиц [20], примеры НКРЯ, а также сравнительный анализ семантически близкого глагола *бахать*, предполагаем, что представленное в словарях количество модификаций является далеко не полным, потенциальных модификаций может быть больше. Предположение основано на дефиниционном анализе исходных глаголов *бахать* и *бабахать*. Сравним их значения (здесь и далее толкования глаголов приводятся в сокращенном виде).

Бахать. 1. Издавать сильный низкий звук (обычно о пушках и минометах и т. п.). 2. Раздаваться, слышаться [15].

Бабахать. 1. Издавать сильные отрывистые звуки или раздаваться (о звуке); издавать сильные отрывистые звуки, шум, грохот от выстрела, разрыва, падения (обычно о пушках, минометах и т. п.). 2. С силой ударять, стучать, выстреливать [Там же].

Значение *бабахать* на первый взгляд шире (есть дополнительное ЛСВ – «с силой ударять, выстреливать»), но модификационное гнездо у данного глагола существенно меньше, чем у синонима *бахать*. Ср., например:

Состав модификационного гнезда бахать: бахнуть, забахать, отбахать, побахать, пробахать₁, пробахать₂, разбахаться;

Состав модификационного гнезда бабахать: бабахнуть, забаббахать.

Видим, что лакунообразующими модификаторами для глагола *бабахать* выступают: *от-*, *по-*, *про1-*, *про2-*, *раз-*:

Mod4 *бахнуть* ----> Mod4 *бабахнуть*

Mod2 *забахать* ----> Mod2 *забабахать*

Mod9 *отбахать* ----> нет модификации Mod9 *отбабахать*

Mod1 *побахать* ----> нет модификации Mod1 *побабахать*

Mod1 *пробахать₁* ----> нет модификации Mod1 *пробабахать₁*

Mod3 *пробахать₂* ----> нет модификации Mod3 *пробабахать₂*

Mod5 *разбахаться* ----> нет модификации Mod5 *разбабахаться*

Данные поисковой программы в НКРЯ показали еще одну модификацию – *побабахать*, которую мы относим к потенциальным:

ПОБАБАХАТЬ. *Пока же мы пользуемся ... как дикари в набедренной повязке – компьютером: можно и орех расколотить, и спину почесать, и по кнопкам побаббахать, от надоедливой соседа отмахаться...* (Лариса Шульман. Сквозняки будущего Штрихи к жизни и литературе) [13].

У глаголов *бряцать* найдены модификации Mod4 *бряцнуть* и Mod1 *побряцать* (в двух значениях: прямом звуковом и переносном):

Mod4 бряцнуть: *Где-то еле слышно бряцнул колокол. «Два», – машинально отметил я (В. Скрипкин. Тинга) [Там же]. – Гоша поддел брезентовый мешок под столом, тот бряцнул в ответ (Ян Разливинский. Старьевщик) [Там же].*

Mod1 побряцать: *Их нельзя будет потерять, нечем будет побряцать в кармане (Дмитрий Урусиков. Я буду жить в виртуальном мире) [Там же].*

Mod1 побряцать (переносное): *А в мирное время полетал-полетал, побряцал оружием и летишь домой на авианосец (Михаил Панин. Камикадзе) [Там же]. Эти голоса по-иному толковали все происшествие: кто-то, мол, побряцал увесистой мошной, чтобы обратить в прах лондонских инспекторов... (Роберт Штильмарк. Наследник из Калькутты) [Там же].*

Все переносные значения выводятся из НКРЯ и других источников вручную, поисковая система по понятным причинам не в состоянии пока распознавать данные семантические нюансы.

На последнем этапе поиска потенциальных модификаций выстраиваем список потенциальных лексем по группам образующих модификаторов:

Mod1 ПО-, ПРО₁: *побахать, побренькать, поверещать, погикать, погорланить; пожурчать, поквакать, покукарекать, помекать, помяукать, порыдать, попищать, порыкать, порявкать, посипеть, постенать, потенькать, потикать, потопотать, потукать, поурчать, поухать, похлипать, почиркать, пошелестеть; пробахать, пробрехать, пробрякать, прогалдеть, прогаркать, прогикать, прогукать, прозвякать, прокурлыкать, промекать, простенать, проухать, прохлюпать, прочавкать, прочивикать;*

Mod9 ОТ-: *отбахать, отчивикать;*

Mod6 ПО-ИВА-: *поахивать; покликивать, покурлыкивать, порявкивать, похрупывать, похрястывать, похряпывать;*

Mod5 РАЗ-/РАС-...СЯ: *разбахаться, разбрехаться, развякаться, разгикаться, раздудеться, размекаться, размычаться, разрыкаться, разрычаться, разухаться, разыкаться, распиликатся, расстучаться, растопаться, растопотаться, растукаться, расчивикаться, расшипеться;*

Mod3 ПРО₂-, С-: *скричать и т. п.*

Эти и другие потенциальные модификации, выявленные с помощью метода веб-извлечения, могут быть лексикографированы как в толковых, так и в аспектных словарях различной направленности (морфемных, синонимических, модификационных и т. п.).

Заключение

Проведенное исследование всех типов модификационной лексики глаголов звучания позволяет сделать следующие выводы.

1. Для систематизации терминологии, относящейся к категории глагольного синтеза, введенные термины «потенциальные модификации» и «лакунообразующие модификаторы» наиболее точно определяют новые лексические единицы, не зафиксированные словарями, но способные, с одной стороны, отражать процессы синтеза, с другой – образовывать лакуны – невозможные сочетания в языке. Анализ потенциальной лексики и лакунообразующих модификаторов группы глаголов звучания показывает, что каждый модификатор, в зависимости от исходного глагола, может быть как активным (приводить к образованию узуальной лексики), так и «нулевым» – лакунообразующим.

2. Условные лакуны, или потенциальные модификации, относятся к категории лексем, которые формально не существуют. Но относить их к разряду лакун не следует. Потенциальные модификации образуют особую лексическую группу, граничащую с лакунами лишь формально, по признаку некодифицированности. Главное отличие их от лакун в том, что потенциальные модификации могут изменить свой статус и войти в состав узуальной лексики. Следует лишь доказать, что данные слова употребляются и фиксируются в НКРЯ и других источниках.

3. Анализ значений потенциальных единиц на уровне ЛСВ показывает, что дифференциация и индексация вариантов контекстных значений позволяет выявить максимально верное количество синтезированных лексем и количество потенциальных «звуковых» модификаций.

3. Методика веб-извлечения потенциальных единиц синтеза модификаций позволяет ответить на ряд вопросов, которые раньше решались вручную и

интуитивно: выявлять наиболее частотные модификаторы, наиболее продуктивные глаголы, выводить из состава лакун те модификации, которые были к ним причислены из-за отсутствия в словарях.

Перспективы дальнейшего исследования мы видим в изучении категории потенциальных модификаций на материале других лексико-семантических групп и построения словарей модификаторов и модификаций.

Список литературы

1. Бабенко Л.Г. Моделирование семантики русских синонимов в идеографическом словаре (на материале словарных статей синонимов с эмотивной семантикой) / Л.Г. Бабенко // Вопросы лексикографии. Russian journal of lexicography. – Томск, 2018. – №13. – С. 81–100.
2. Земская Е.А. Производные слова в толковых словарях русского языка / Е.А. Земская // Современная русская лексикография. – Л., 1977.
3. Ивлиева И.В. Особенности синтеза многозначных глаголов. Семантическое варьирование морфем / И.В. Ивлиева // Вопросы филологических наук. – 2003. – №4.
4. Ивлиева И.В. Проблемы синтеза русского глагола (на материале антропофонов): монография / И.В. Ивлиева. – М: Изд-во РУДН, 2003. – 120 с.
5. Ивлиева И.В. Русские глагольные модификации (опыт составления словаря) / И.В. Ивлиева. – М., 2008. – 275 с.
6. Ивлиева И.В. Экспериментальный модификационный словарь русского языка (на материале глаголов звучания) / И.В. Ивлиева. – М., 2013. – 467 с.
7. Ивлиева И.В. Семантика «лакунных модификаций» (на материале глаголов звучания) / И.В. Ивлиева // Интерактивная наука. – 2016. – №4 (26).
8. Катлинская Л.П. Проблемы слова и словообразования / Л.П. Катлинская // Филологические науки. – 2003. – №3. – С.102.
9. Кронгауз М.А. Приставки и глаголы / М.А. Кронгауз. – М., 1998.
10. Кубрякова Е.С. Семантика производного слова / Е.С. Кубрякова // Аспекты семантических исследований. – М.: Наука, 1980. – С. 81–155.

11. Ляшевская О.Н. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка) / О.Н. Ляшевская, С.А. Шаров. – М.: Азбуковник, 2009.
12. Милославский И.Г. Вопросы словообразовательного синтеза / И.Г. Милославский. – М., 1980.
13. Национальный корпус русского языка [Электронный ресурс]. – Режим доступа: <http://www.ruscorpora.ru/>
14. Сводный словарь современной русской лексики: В 2 т. / АН СССР. Ин-т рус. яз.; под ред. Р.П. Рогожниковой. – М.: Рус. яз., 1990. – Т. 1. – 800 с.
15. Современный толковый словарь русского языка Т.Ф. Ефремовой [Электронный ресурс]. – Режим доступа: <https://dic.academic.ru/contents.nsf/efremova/>
16. Тихонов А.Н. О семантической соотносительности производных и производящих основ / А.Н. Тихонов // Вопросы языкознания. – 1967. – №1. – С. 112 – 120.
17. Торопцев И.С. Словопроизводственная модель / И.С. Торопцев. – Воронеж: Изд-во Воронежского ун-та, 1980. – 148 с.
18. Улуханов И.С. единицы словообразовательной системы русского языка и их лексическая реализация / И.С. Улуханов. – М., 1996. – 221 с.
19. Янко-Триницкая Н.А. Закономерность связей словообразовательного и лексического значений в производных словах / Н.А. Янко-Триницкая // Развитие современного русского языка. – М.: Наука, 1963. – С. 83 – 98.
20. Koob, P., Ivliyeva, I. An Assessment of the Propagation of Contemporary Russian Mechanical Phonation Verbs to Online Dictionaries. Web scraping project by Missouri S&T, IT and ALP departments. Retrieved from: https://scholarsmine.mst.edu/artlan_phil_facwork/157/
21. Koob, P., Ivliyeva, I. Russian National Corpus Web Scraping Project 2019-2020. Retrieved from: https://scholarsmine.mst.edu/artlan_phil_facwork/158/

Ивлиева Ирина Валентиновна – канд. филол. наук, профессор кафедры искусств, иностранных языков и философии, Научно-технологический университет штата Миссури, Ролла, США.
